Available online at www.sciencedirect.com
ScienceDirect
Procedia Engineering 154 (2016) 36 – 41
Procedia Engineering
www.elsevier.com/locate/procedia

12th International Conference on Hydroinformatics, HIC 2016

Automated detection of faults in wastewater pipes from CCTV footage by using Random Forests

Joshua Myrans[a], Zoran Kapelan[b] & Richard Everson[c]

[a]*PhD student, University of Exeter, Harrison Building, North Park Road, Exeter, Devon, UK*
[b]*Professor of Water Systems Engineeing, University of Exeter, Harrison Building, North Park Road, Exeter, Devon, UK*
[c]*Professor of Machine Learning, University of Exeter, Harrison Building, North Park Road, Exeter, Devon, UK*

Abstract

Sewer systems require regular inspection in order to ensure their satisfactory condition. As most sewer networks consist of pipes too small for engineers to traverse, CCTV footage is used to record the interior of these pipes. This footage is manually analysed by qualified engineers, to determine the condition of the pipe and the presence of any faults. We propose a methodology, which automatically detects faults within the CCTV footage. This has the potential to dramatically reduce the time required to process the large volume of CCTV footage produced during a survey. The proposed methodology first characterises localised regions of each video frame using multiscale GIST features. Extremely randomised trees are then used to learn a classifier that distinguishes between frames showing a fault and normal frames. The technique is tested on 670 video segments from real sewer inspections of a variety of pipes, supplied by Wessex Water. Detection performance is assessed by plotting receiver operating characteristics and quantifying the area under the curve. Preliminary results indicate high detection accuracy of 88% and an area under the ROC curve of 96%. The machine learning used reduces the footage to a selection of frames containing faults, which can be quickly identified (whether by an engineer or another piece of software), showing promise for use in industrial wastewater network surveys.

Peer-review under responsibility of the organizing committee of HIC 2016

Keywords: Automation; Fault detection; Sewers; GIST; Random Forests; Extra Trees;

Peer-review under responsibility of the organizing committee of HIC 2016
doi:10.1016/j.proeng.2016.07.416

1. Introduction

Wastewater networks around the world require regular inspection in order to prioritise and perform effective maintenance. Currently wastewater networks are inspected using CCTV, taking one of two approaches. The first requires a camera, attached to a semi-rigid wire, to be pushed through the network. In doing so the collected footage is left to be analysed later by a trained engineer. Although quick to collect, the footage gathered is often of lower quality as the camera does not travel smoothly through the pipe. Alternatively a camera can be attached to a remote controlled PIG (pipe inspection gadget) which is driven through the network. Due to the live feed and superior control provided by the PIG, a skilled operator can often identify and record faults whilst operating the device. In doing so footage takes longer to collect, but does not require further analysis and is often of higher quality.

Both approaches are expensive, requiring a trained engineer for the duration of the survey and the follow-on analyses. Collecting footage using a wire is cheap, although the later analysis is slow and expensive, whilst using a PIG is slow and expensive, especially as PIG operation requires additional training. Additionally, both methods rely on the training and experience of the operator/analyst, even given a formal guideline [1], operators are only human. It is not uncommon for a single fault to be classified incorrectly or overlooked, due to the subjective nature of most faults [2].

The proposed methodology attempts to counter the problems identified above. The method aims to classify frames of CCTV footage into faulty and normal categories, which can be further categorised by an engineer. By doing so the methodology attempts to reduce days and hours of footage down to a selection of relevant frames, most of which contain only faults. It is hoped that this will in turn reduce the costs (both time and money) associated with surveys, whilst preserving (or improving) a survey's accuracy.

2. Background

A limited amount of work has been previously undertaken in this field, most notably by Duran et al [3] and Sinha et al [4]. Duran chose to retrofit the traditional CCTV camera with a laser profiler, in order to get accurate information about the pipe's interior. Passing the profiler's readings through a series of ANNs (artificial neural networks), Duran was able to accurately (90% +) classify a selection of faults fabricated in a laboratory. On the other hand, Sinha applied fuzzy logic to the problem, identifying five characteristics, which could be reasonably measured and were key to detecting a fault. These were light intensity, texture, size (major and minor axis lengths), shape and organisation. The features, once fuzzified, were again passed to an ANN classifier, and produced accuracies of 85-95%, for joints, cracks and connections on flush cleaned concrete pipes.

Neither of these approaches was demonstrated using raw footage from real CCTV surveys Duran's methodology was only tested in laboratory-based, fabricated experiments, whilst Sinha's methodology was tested on flush cleaned pipes. This leaves both methodologies untested in a 'real world' environment, and their usefulness to industry unproven. Furthermore, both Duran and Sinha choose to use ANNs (Artificial Neural Networks), a black box approach to classification. In opting to do so, information about a classification is lost, providing little justification for each decision. This lack of understanding and accountability makes it harder to justify detections and classifications made by Duran and Sinha's methodologies.

3. Fault Detection Methodology

Acknowledging previous work in the field, this methodology aims to work from actual survey data, in order to ensure its applicability to 'real life' systems. The methodology was developed on footage taken from actual surveys undertaken by the Wessex Water, collected using both wire driven cameras and PIGs. This footage was processed, before being classified by a trained random forest in order to identify whether individual frames contained faults. In order to prepare the random forest, a separate library of processed frames was labelled. The labelled frames were then used to train the classifier, following the extra trees methodology [7]. This developed methodology can be broken down into three general steps:

1. Frame extraction and pre-processing
2. Descriptor calculation
3. Detection.

Following the training of the classifier these steps individual frames can be lifted from CCTV footage, and the contents identified as either normal or faulty. By applying the outlined steps to sequences of frames extracted from a survey, the methodology should classify each frame, allowing for easy identification of faults within the footage.

3.1. Frame Extraction and Processing

The initial step of the methodology requires frames to be extracted from the raw CCTV footage, the rate of which is dependent upon the traveling speed of the camera. This could be as infrequent as once per second, as cameras often travel at slow rate in order to collect higher quality images [1]. An extracted frame is then rescaled, normalised and converted to a grey-scale image. Normalisation is performed locally, by dividing by the variance in local pixel intensity, in order to limit the influence of the light source's orientation. Processing each frame before feature extraction, ensures each image has uniform properties, making it possible to compare frames across different surveys, which is key in training the classifier.

3.2. Image Descriptor Calculation

Given a processed frame, its GIST [5] feature descriptor can be calculated. A GIST feature attempts to represent the overall contents of the image, looking at large patches of the image in order to generate a holistic descriptor. Calculating a GIST descriptor requires convolving the frame with a sequence of Gabor wavelets. The wavelets are transformed to cover 8 orientations across 4 scales, as shown in Figure 1. These values were selected after preliminary sensitivity analyses. Each convolution produces a feature map, for a total of 32 maps. Every map is then split using the same four by four grid into 16 cells, as shown in Figure 2. Finally the contents of each cell is totalled and appended to a composite feature vector. The result is a 512-dimensional vector containing the sum of each of the 16 cells for each of the 32 feature maps [5].

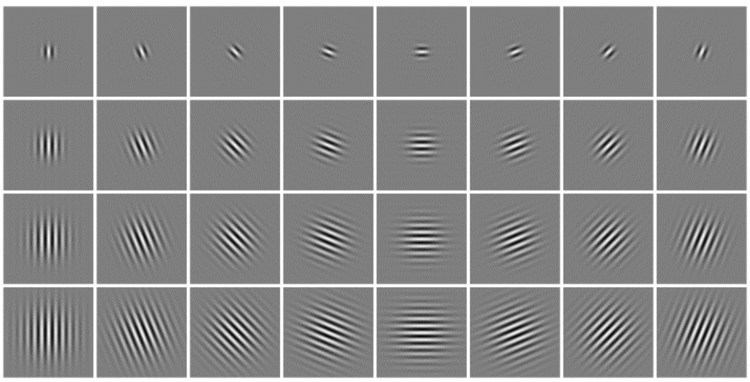

Fig. 1: The 4 scales and 8 orientations of Gabor wavelets used in the calculation of GIST features.

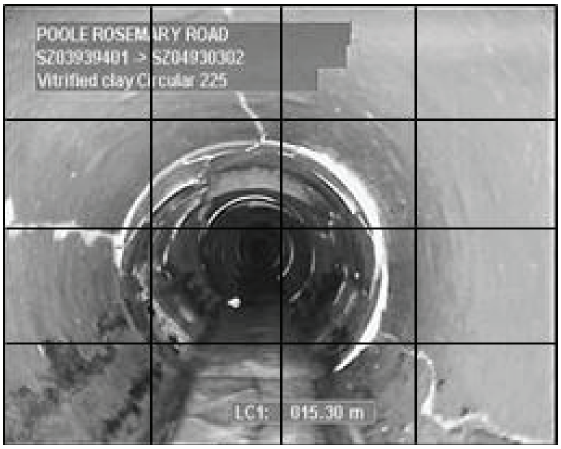

Fig. 2: An example frame, demonstrating the overlaid 4x4 grid

3.3. Detection

Having characterised frames as GIST feature vectors, a trained Random Forest (RF) classifier [6] is used here. Random Forests work by training an ensemble of individual decision trees, on a library of labelled training data, which has been processed using the above steps. The library is used to train each of the trees in the forest, using the Extra Trees algorithm [7]. The Extra Trees algorithm, emphasises the randomness of each individual tree, randomising both feature and threshold on which a tree's decision is made. Prepared, the random forest can make accurate predictions on members of the training library, and extrapolates to unseen feature vectors. It should be noted that as with all classifiers, techniques such as cross validation should be implemented in order to avoid over/under fitting the classifier.

Once trained, the forest is then passed unseen feature vectors (generated from extracted frames) whose label (faulty or normal is unknown. Each tree in the forest independently classifies the given feature vector, identifying the vector as either faulty or normal. After all trees have made a decision, each tree votes on the vector's identity, with the prevailing identity being selected as the vector's classification.

Random Forests were chosen for this methodology, due to their relative transparency in comparison to other classifiers of similar performance. Each identification performed by a forest can be traced back through every tree, where the decisions leading to an identification can be observed, and even located within the original image thanks to the use of GIST descriptors. This property of Random Forests, provides an insight into how classifications are made, providing justification for every frame's identification, which could be useful in further locating and classifying fault types, whilst reassuring the users of the classifiers decision.

4. Case Study

The methodology was implemented and tested on the CCTV footage provided by Wessex Water, which was taken directly from surveys used by their engineers. A library of frames was extracted from 5 separate surveys covering over 5.5 km of wastewater pipes with diameters ranging from 150 to 1800 mm. These pipes varied in both shape and material, containing a mix of vitrified clay, brick and metal pipes of various shapes. From the 15 hours of CCTV footage, 670 frames were manually extracted and processed as per the methodology, 312 of which contained faults of various types. These faults were extracted using the surveyors' annotations that accompanied the footage, ensuring their accuracy and are described in Table 1. The remaining 358 images were extracted at random from all surveys, and manually checked to ensure they contained no faults.

Table 1: Distribution of fault types in the library of extracted frames.

Fault type	Subtypes	Percentage (%)
Joint	Displaced, Open	24
Crack	Longitudinal, Circumferential, Multiple, Spiral	11
Broken / Collapsed	-	3
Obstacles	Intruding junctions, Masonry, Protrusions	15
Hole	-	11
Deposits	Attached, Settled	28
Roots	Fine, Tap, Mass	3
Infiltration	Running, Gushing	1
Brickwork	Missing mortar, Displaced bricks, Missing bricks	5

Following the prescribed steps each frame was processed, and its GIST feature vector calculated. Having a library of GIST features, a random forest of Extra-Trees [7] was trained. In order to avoid overfitting, and have confidence in our results, 25 fold cross validation was performed on the shuffled library [8]. This process yielded an average correct detection rate of 88% on the testing folds. This rate was achieved by arbitrarily setting the classification threshold to 0.5, requiring over half the trees in the random forest to designate a frame as faulty for it to be flagged. In order to better understand the classifiers performance, the ROC curve, shown in Figure 3, was calculated.

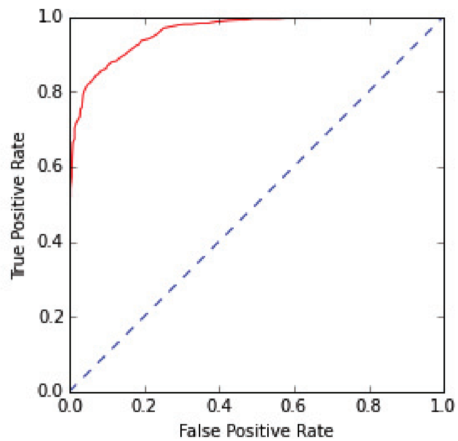

Fig. 3: Receiver operating curve, demonstrating the effect of manipulating the Random Forest's threshold on true and false positive rates

From the ROC curve it is clear to see the effectiveness of the methodology, as the curve pushes towards goal of 100% true positive and 0% false positive rates, further emphasised by an area under the curve of 96%. From this a fault detection threshold can be selected, based on the importance of a pipe. For example, when surveying a pipe that services an entire neighbourhood, the engineer may choose a threshold that results in more false negatives, in order to catch as many faults as possible.

5. Conclusion

This paper has described a novel method of automatically detecting faults in wastewater pipes from CCTV inspection footage. The method harnesses the power of large data sets by training a powerful classifier on entire frames rather than manually defined features, an approach that has been successful in other computer vision applications. The developed methodology, offers a new approach to identification of faults in wastewater networks, and displays promising results for its extension and application in water industry. It has been shown that a GIST feature vector can significantly reduce the dimensionality of footage, whilst retaining the features required to make an accurate prediction about a frame's state. Furthermore Random Forests have proved effective classifiers, capable of correctly distinguishing between 88% of the frames, whilst remaining interpretable.

In conjunction the described techniques and overall methodology could lead to the production of a tool capable of significantly improving the speed of surveying a wastewater network. In order to achieve such a goal, additional survey footage needs to be analysed first. Even with the diverse selection of faults, pipes and materials used in the case study here, this still isn't representative enough of the thousands of kilometres of wastewater network in the UK alone.

6. Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council in the UK via grant EP/L0116214/1 awarded for the Water Informatics, Science and Engineering (WISE) centre for doctoral training which is gratefully acknowledged.

The data utilised in our case study, was kindly provided by Wessex Water, who we gratefully acknowledge. Moreover we would like to thank Julian Britain (Wessex Water) who gave up his time to provide the raw CCTV footage and an insight into the collection of the footage.

7. References

[1] WRc, 2013. Manual of Sewer Condition Classification. 5th Revised edition. WRc Publications.
[2] Rahman, S. and Vanier, D.J., 2004. An evaluation of condition assessment protocols for sewer management.
[3] Duran, O., Althoefer, K. and Seneviratne, L.D., 2007. Automated pipe defect detection and categorization using camera/laser-based profiler and artificial neural network. Automation Science and Engineering, IEEE Transactions on, 4(1), pp.118-126.
[4] Sinha, S.K. and Fieguth, P.W., 2006. Neuro-fuzzy network for the classification of buried pipe defects. Automation in Construction, 15(1), pp.73-83.
[5] Oliva, A. and Torralba, A., 2001. Modelling the shape of the scene: A holistic representation of the spatial envelope. International Journal of Computer Vision, 42(3), pp.145-175.
[6] Breiman, L., 2001. Random Forests. Machine learning, 45(1), pp.5-32.
[7] Geurts, P., Ernst, D. and Wehenkel, L., 2006. Extremely randomized trees. Machine learning, 63(1), pp.3-42.
[8] Kohavi, R., 1995, August. A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai (Vol. 14, No. 2, pp. 1137-1145).